Numer Algor (2014) 67:385–403
DOI 10.1007/s11075-013-9798-4

ORIGINAL PAPER

Efficient algorithms for multivariate and ∞-variate integration with exponential weight

L. Plaskota · G. W. Wasilkowski

Received: 13 June 2013 / Accepted: 3 November 2013 / Published online: 16 November 2013

Abstract Using the *Multivariate Decomposition Method (MDM)*, we develop an efficient algorithm for approximating the ∞-variate integral

$$\mathcal{I}_\infty(f) = \lim_{d\to\infty} \int_{\mathbb{R}_+^d} f(x_1, \ldots, x_d, 0, 0, \ldots) \cdot \exp\left(-\sum_{j=1}^{d} x_j\right) \mathrm{d}\boldsymbol{x}$$

for a class of functions f that are once differentiable with respect to each variable. MDM requires efficient algorithms for d-variate versions of the problem. Such algorithms are provided by Smolyak's construction which is based on efficient algorithms for the univariate integration

$$I(f) = \int_0^\infty f(x)\, \mathrm{e}^{-x}\, \mathrm{d}x.$$

Detailed analysis and development of (nearly) optimal quadratures for $I(f)$ is the main contribution of the current paper.

Keywords Multivariate integration · Weighted integrals · Weighted trapezoid rules · ∞-variate functions

L. Plaskota (✉)
Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, ul. Banacha 2, 02-097 Warsaw, Poland
e-mail: leszekp@mimuw.edu.pl

G. W. Wasilkowski
Department of Computer Science, University of Kentucky, 329 Rose Street, Lexington, KY 40506, USA
e-mail: greg@cs.uky.edu

1 Introduction

In this paper, we study efficient algorithms for approximating multivariate and ∞-variate weighted integrals with the exponential probability density weight. That is, we are interested in approximating

$$I_d(f) = \int_{\mathbb{R}_+^d} f(\boldsymbol{x}) \cdot \exp\left(-\sum_{j=1}^{d} x_j\right) \mathrm{d}\boldsymbol{x}$$

and its extension to integrals of functions with countably many variables.

The algorithms use evaluations of f at points with only finitely many nonzero coordinates. The cost of obtaining $f(\boldsymbol{x})$ equals $\$(d)$ for a cost function $\$$, where d is the number of nonzero coordinates in $\boldsymbol{x}$. The cost of an algorithm is the total cost of such evaluations.

Theoretical results obtained here will be used in the future, as we plan to develop and test algorithms for different types of probability density functions including the exponential one. There is a growing interest in the complexity/tractability of ∞-variate integration and/or approximation, see, e.g., [1–11, 13–16, 19–23, 25–27]. The current paper is an addition to this list.

In the d-variate case, the integrands belong to the space F_d of functions that vanish at $\boldsymbol{x} = (x_1, \ldots, x_d)$ if at least one $x_j = 0$. Moreover, the norm $\|f\|_{F_d}$ is given by the L_∞ norm of mixed first order partial derivatives of f.

In the ∞-variate case, the integrands are from the following space $\mathcal{F} = \mathcal{F}_{\gamma,q}$. For a finite subset $\mathfrak{u} = \{u_1, \ldots, u_d\}$ of the set $\mathbb{N}$ of positive integers, let $F_\mathfrak{u}$ be like the space F_d for $d = |\mathfrak{u}|$ with the only difference that the variables of functions from $F_\mathfrak{u}$ are $x_{u_1}, \ldots, x_{u_d}$ instead of $x_1, \ldots, x_d$. Then $\mathcal{F}_{\gamma,q}$ is the completion of the span of $\bigcup_{|\mathfrak{u}|<\infty} F_\mathfrak{u}$ with respect to the norm

$$\|f\|_{\mathcal{F}_{\gamma,q}} = \left[(|f_\emptyset|/\gamma_\emptyset)^q + \sum_{|\mathfrak{u}|<\infty} (\|f_\mathfrak{u}\|_{F_\mathfrak{u}}/\gamma_\mathfrak{u})^q \right]^{1/q}.$$

Here $q \in [1, \infty]$ and $\gamma = \{\gamma_\mathfrak{u}\}_\mathfrak{u}$ is a family of nonnegative numbers called *weights*. Observe that, due to the vanishing aspect of $f_\mathfrak{u}(\boldsymbol{x})$ for $\boldsymbol{x}$ with some zero components, any function $f \in \mathcal{F}_{\gamma,q}$ has a unique decomposition

$$f = \sum_{|\mathfrak{u}|<\infty} f_\mathfrak{u} \quad \text{with} \quad f_\mathfrak{u} \in F_\mathfrak{u}. \tag{1}$$

Recall that an ∞-variate problem is *polynomially tractable* if the minimal cost of computing an ε-approximation is bounded from above by $C\,\varepsilon^{-p}$ for any $\varepsilon \in (0, 1)$. The minimal such exponent p is called the *exponent of polynomial tractability* and denoted by p^{trct}. The problem is *weakly tractable* if this minimal cost is **not** exponential in $1/\varepsilon$.

We provide conditions on γ, q and $\$$ for polynomial and weak tractabilities. We also provide the exact value of p^{trct} for some special cases. In particular, assume that the weights have the *product order-dependent* (*POD* for short) form

$$\gamma_{\mathfrak{u}} = (|\mathfrak{u}|!)^{\beta_1} \cdot \prod_{j \in \mathfrak{u}} \frac{c}{j^{\beta_2}} \quad \text{with} \quad \beta_2 > \max(0, \beta_1) \quad \text{and} \quad c > 0. \tag{2}$$

If $\beta_2 > 1 - 1/q$, then the tractability exponent is given by

$$p^{\mathrm{trct}} = \max\left(1\,,\, \frac{1}{\beta_2 - 1 + 1/q}\right).$$

We stress that polynomial tractability is possible even if the cost function \$ is exponential in d. Weak tractability is possible even if \$ is doubly exponential in d. Moreover, the results are constructive.

Indeed, following [22] we apply the *Multivariate Decomposition Method* (*MDM* for short) for our ∞-variate integration. This method evolved from the Changing Dimension Algorithm introduced in [11]. It requires efficient algorithms for the d-variate versions of the problem. Such algorithms are derived using Smolyak's construction [17] and the results of [24]. Smolyak's construction requires efficient algorithms for the univariate integration

$$I(f) = \int_0^\infty f(x)\, \mathrm{e}^{-x}\, \mathrm{d}x.$$

Detailed analysis and development of (nearly) optimal quadratures for the one dimensional integration is the main contribution of the current paper. We propose to use a weighted version of (composite) trapezoidal rules T_n^* at points

$$x_{n,i}^* = -2\, \ln\left(1 - \frac{i}{n+1}\right) \quad \text{for} \quad i = 0, 1, \ldots, n.$$

More precisely, the value of the quadrature is the integral of the piecewise linear function p interpolating f at the points $x_{n,i}^*$ and such that $p(x) = f(x_{n,n}^*)$ for $x \geq x_{n,n}^*$. We show that the worst case error of T_n^*, i.e., $\sup_{\|f\|_{F_1} \leq 1} |I(f) - T_n^*(f)| = \|I - T_n^*\|$, satisfies

$$\|I - T_n^*\|\,(n+1) = (1 + O(1/n)).$$

Hence T_n^* is nearly optimal, as we also prove that any algorithm using n points has the worst case error larger than $(n+1)^{-1}$ in the space F_1.

The error bounds of Smolyak's algorithm for d-variate integration depend on

$$C_1 := \sup_{j \geq 1} \|I - T_{2^j-1}^*\| \cdot 2^j$$

and $C = \max(C_0/2, C_2)$, where $C_0 = \|I\| = 1$ and

$$C_2 := \sup_{j \geq 1} \|T_{2^j-1}^* - T_{2^{j-1}-1}^*\| \cdot 2^j.$$

The dependence on C is essential since the error bounds of Smolyak's algorithm are proportional to $(C/2)^{d-1}$. We prove that

$$C_1 = 2\,\|I - T_1^*\| = 1.00655\ldots \qquad \text{and} \qquad C_2 = 2,$$

so that $C/2 = 1$. The derivations of these constants combine analytic estimations with some numerical computations.

We now comment on the choice of the L_∞ norm and the first order derivatives for the definition of the class of functions. The main reason was to have a relatively large class of possible applications. Indeed, for unbounded domains, there are interesting functions with f' bounded in L_∞ norm, but not in L_p for any $p \in (1, \infty)$. This includes $f(x) = (x+1)^a$ for $a \leq 1$. As for the 1st order derivatives (as opposed to higher order derivatives), consider approximating the integral of

$$f(\boldsymbol{x}) = g\left(\sum_{j=1}^{\infty} x_j \, \psi_j\right)$$

for a smooth univariate function g and positive numbers ψ_j that converge to zero sufficiently fast. Even though we do not know the specific terms $f_{\mathfrak{u}}$ of the decomposition (1), we have

$$\|f_{\mathfrak{u}}\|_{F_{\mathfrak{u}}} \leq \left\| g^{(|\mathfrak{u}|)} \right\|_\infty \cdot \prod_{j \in \mathfrak{u}} \psi_j .$$

Therefore, for specific functions g, it might be possible to estimate the norms of $f_{\mathfrak{u}}$ and choose appropriate weights $\gamma_{\mathfrak{u}}$; see [9, 10] where this idea has been used for approximating unweighted integrals of the form

$$\mathcal{I}_\infty(f) = \int_{[0,1]^{\mathbb{N}}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}.$$

Although there are results and various quadratures for weighted univariate integration, including Gaussian rules, we stress that none of them addresses the worst case complexity and (nearly) optimal quadratures for the space of functions with bounded $\|f'\|_{L_\infty(\mathbb{R}_+)}$. As mentioned above, this space is of significant importance for problems involving ∞-variate integrals.

The paper is organized as follows. The weighted trapezoidal quadrature for the univariate case is presented and the constants C_1 and C_2 are derived in Section 2. The Smolyak's construction for the d-variate case is in Section 3. Section 4 is devoted to MDM for the ∞-variate case.

2 Quadratures for univariate weighted integrals

We consider a weighted integration of scalar functions $f : [0, \infty) \to \mathbb{R}$,

$$I(f) = \int_0^\infty f(x) \, \rho(x) \, \mathrm{d}x,$$

where ρ is the exponential weight

$$\rho(x) = \mathrm{e}^{-x}. \tag{3}$$

We assume that the integrands belong to the space F of functions f such that $f(0) = 0$, f is absolutely continuous on any subinterval $[0, T]$, $0 < T < \infty$, and the norm

$$\|f\|_F := \operatorname*{ess\,sup}_{0 \leq x < \infty} |f'(x)| < \infty.$$

Note that

$$|I(f)| = \left|\int_0^\infty f'(t)\int_t^\infty \mathrm{e}^{-x}\,\mathrm{d}x\,\mathrm{d}t\right| \le \|f\|_F \int_0^\infty \mathrm{e}^{-t}\,\mathrm{d}t = \|f\|_F.$$

Since this inequality is sharp, the functional I is continuous and its norm

$$\|I\| = \sup_{\|f\|_F\le 1} |I(f)| = 1. \tag{4}$$

The integral $I(f)$ is approximated by the ρ-weighted composite trapezoidal rule $T_n(f)$. That is, $T_n(f) = I(p_f)$, where p_f is the piecewise linear function that interpolates f at points x_i,

$$0 = x_0 < x_1 < \cdots < x_n < +\infty,$$

with $p_f(t) \equiv f(x_n)$ for $t \ge x_n$. In the explicit form,

$$T_n(f) = \sum_{i=1}^n a_i\, f(x_i),$$

where

$$a_i = \frac{\mathrm{e}^{-x_{i+1}} - \mathrm{e}^{-x_i}}{x_{i+1}-x_i} - \frac{\mathrm{e}^{-x_i} - \mathrm{e}^{-x_{i-1}}}{x_i - x_{i-1}}, \qquad 1 \le i \le n-1,$$

and

$$a_n = -\frac{\mathrm{e}^{-x_n} - \mathrm{e}^{-x_{n-1}}}{x_n - x_{n-1}}.$$

We now find a formula for the error $I(f) - T_n(f)$. The 'local' error for each of the subintervals $[x_{i-1}, x_i]$, $1 \le i \le n$, equals

$$\begin{aligned}
&\int_{x_{i-1}}^{x_i} (f(x) - p_f(x))\cdot\rho(x)\,\mathrm{d}x \\
&= \int_{x_{i-1}}^{x_i}\left(\int_t^{x_i}\rho(x)\,\mathrm{d}x - \int_{x_{i-1}}^{x_i}\frac{x - x_{i-1}}{x_i - x_{i-1}}\cdot\rho(x)\,\mathrm{d}x\right) f'(t)\,\mathrm{d}t \\
&= \int_{x_{i-1}}^{x_i}\left(\int_{x_{i-1}}^{x_i} B_{n,i}(x,t)\cdot\rho(x)\,\mathrm{d}x\right) f'(t)\,\mathrm{d}t,
\end{aligned}$$

where

$$B_{n,i}(x,t) = \mathbf{1}_{[t,x_i]}(x) - \frac{x - x_{i-1}}{x_i - x_{i-1}}.$$

Here and elsewhere, $\mathbf{1}_D$ is the characteristic function of a set $D \subset \mathbb{R}$. Similarly,

$$\int_{x_n}^\infty (f(x) - p_f(x))\cdot\rho(x)\,\mathrm{d}x = \int_{x_n}^\infty\left(\int_{x_n}^\infty B_{n,n+1}(x,t)\cdot\rho(x)\,\mathrm{d}x\right) f'(t)\,\mathrm{d}t,$$

where

$$B_{n,n+1}(x,t) = \mathbf{1}_{[t,\infty)}(x).$$

Summarizing, we have

$$I(f) - T_n(f) = \int_0^\infty K_n(t)\, f'(t)\,\mathrm{d}t, \tag{5}$$

with the Peano kernel

$$K_n(t) \;=\; \int_{x_{i-1}}^{x_i} B_{n,i}(x,t)\,\rho(x)\,\mathrm{d}x, \qquad x_{i-1} \le t < x_i.$$

($x_{n+1} = \infty$ by convention.)

For the exponential weight (3), these formulas take the following form. For $x_{i-1} \le t < x_i$, $1 \le i \le n$, we have

$$\begin{aligned} K_n(t) \;&=\; \int_{x_{i-1}}^{x_i} B_{n,i}(x,t)\,\mathrm{e}^{-x}\,\mathrm{d}x \\ &=\; -\int_{x_{i-1}}^{t} \frac{x - x_{i-1}}{x_i - x_{i-1}} \cdot \mathrm{e}^{-x}\,\mathrm{d}x + \int_{t}^{x_i} \frac{x_i - x}{x_i - x_{i-1}} \cdot \mathrm{e}^{-x}\,\mathrm{d}x \\ &=\; -\frac{\mathrm{e}^{-x_{i-1}} - \mathrm{e}^{-t}}{x_i - x_{i-1}} + \frac{t - x_{i-1}}{x_i - x_{i-1}} \cdot \mathrm{e}^{-t} + \frac{\mathrm{e}^{-x_i} - \mathrm{e}^{-t}}{x_i - x_{i-1}} + \frac{x_i - t}{x_i - x_{i-1}} \cdot \mathrm{e}^{-t} \\ &=\; \frac{\mathrm{e}^{-x_i} - \mathrm{e}^{-x_{i-1}}}{x_i - x_{i-1}} + \mathrm{e}^{-t}, \end{aligned} \tag{6}$$

and for $x_n \le t < \infty$

$$K_n(t) \;=\; \int_{x_n}^{\infty} B_{n,n+1}(x,t)\,\mathrm{e}^{-x}\,\mathrm{d}x \;=\; \mathrm{e}^{-t}. \tag{7}$$

We now analyze the worst case error of T_n in the space F. Recall that for any algorithm A_n, this error is defined by

$$e(A_n, F) \;:=\; \sup_{\|f\|_F \le 1} |I(f) - A_n(f)|.$$

Since T_n depends linearly on f, we obviously have $e(T_n, F) = \|I - T_n\|$. Next, by (5),

$$\|I - T_n\| \;=\; \int_0^{\infty} |K_n(t)|\,\mathrm{d}t.$$

Direct calculations give

$$\int_{x_{i-1}}^{x_i} |K_n(t)|\,\mathrm{d}t \;=\; 2\left(\frac{\mathrm{e}^{-x_{i-1}} + \mathrm{e}^{-x_i}}{2} - \mathrm{e}^{-z_i} + \left(\frac{x_{i-1} + x_i}{2} - z_i\right)\mathrm{e}^{-z_i}\right), \tag{8}$$

where $x_{i-1} < z_i < x_i$ is such that

$$\mathrm{e}^{-z_i} \;=\; \frac{\mathrm{e}^{-x_{i-1}} - e^{-x_i}}{x_i - x_{i-1}},$$

and $\int_{x_n}^{\infty} |K_n(t)|\,\mathrm{d}t = \mathrm{e}^{-x_n}$. Hence

$$\|I - T_n\| \;=\; \mathrm{e}^{-x_n} + 2 \cdot \sum_{i=1}^{n} \left(\frac{\mathrm{e}^{-x_{i-1}} + \mathrm{e}^{-x_i}}{2} - \mathrm{e}^{-z_i} + \left(\frac{x_{i-1} + x_i}{2} - z_i\right)\mathrm{e}^{-z_i}\right).$$

Proposition 1 *For any choice of the points* $\{x_i\}_{i=1}^n$ *we have*

$$e(T_n, F) \;=\; \|I - T_n\| \;>\; 1/(n+1).$$

Even more, the worst case error of any *algorithm A_n which uses n sample points is larger than* $1/(n+1)$.

Proof Using some geometric arguments, we get that the right-hand side of (8) can be estimated from below by the same expression, but with z_i replaced by the center of the interval $[x_{i-1}, x_i]$. Hence for each i,

$$\int_{x_{i-1}}^{x_i} |K_n(t)|\,\mathrm{d}t \;>\; 2\left(\frac{\mathrm{e}^{-x_{i-1}} - \mathrm{e}^{-x_i}}{2} - \mathrm{e}^{-(x_{i-1}+x_i)/2}\right)$$
$$= \left(\mathrm{e}^{-x_{i-1}/2} - \mathrm{e}^{-x_i/2}\right)^2,$$

and consequently

$$\|I - T_n\| \;>\; \mathrm{e}^{-x_n} + \sum_{i=1}^{n}\left(\mathrm{e}^{-x_{i-1}/2} - \mathrm{e}^{-x_n/2}\right)^2.$$

This is minimized by the points $x_i^* = -2\,\ln\left(1 - \frac{i}{n+1}\right)$ with $1 \le i \le n$, in which case $\mathrm{e}^{-x_i^*/2} = 1 - \frac{i}{n+1}$ and

$$\|I - T_n\| \;>\; \frac{1}{(n+1)^2} + \sum_{i=1}^{n}\frac{1}{(n+1)^2} \;=\; \frac{1}{n+1}.$$

Since the ρ-weighted trapezoidal rule T_n is a central algorithm, see, e.g., [18], it minimizes the worst case error among *all* (even nonlinear) algorithms A_n that use given points $\{x_i\}_{i=1}^n$. The proof is complete. □

As in the proof above, consider

$$x_i^* \;=\; x_{n,i}^* \;:=\; -2\,\ln\left(1 - \frac{i}{n+1}\right), \qquad 1 \le i \le n. \tag{9}$$

Denote by T_n^* the weighted trapezoidal rule that uses these special points, with $T_0^* = 0$.

Theorem 1 *We have*

$$\|I - T_n^*\|\,(n+1) \;=\; 1 + O(n^{-1})$$

and

$$\max_{n\ge 0}\|I - T_n^*\|\,(n+1) \;= 2\,\|I - T_1^*\| \;=\; 1.00655\ldots.$$

Hence T_n^* is nearly optimal.

Proof Letting $k = n - i + 1$, we find by direct calculations that

$$2\left(\frac{\mathrm{e}^{-x_{i-1}^*} + \mathrm{e}^{-x_i^*}}{2} - \mathrm{e}^{-z_i^*}\right) \;=\; \frac{A_k}{(n+1)^2}$$

and

$$2\left(\frac{x_{i-1}^* + x_i^*}{2} - z_i^*\right) \mathrm{e}^{-z_i^*} \;=\; \frac{B_k}{(n+1)^2},$$

where

$$\begin{aligned} A_k \;&=\; 2k^2 + 2k + 1 - \frac{2k+1}{\ln(1+1/k)}, \\ B_k \;&=\; \frac{2k+1}{\ln(1+1/k)} \cdot \ln\left(\frac{k+1/2}{k(k+1)\ln(1+1/k)}\right). \end{aligned}$$

Hence, setting $A_0 = 2/3$, $B_0 = 1/3$, we have

$$\|I - T_n^*\| \;=\; (n+1)^{-2} \cdot \sum_{k=0}^{n} (A_k + B_k).$$

We now estimate A_k and B_k from above using the Taylor expansion of $\ln(1+x)$ for $0 \le x \le 1$. For A_k we use $\ln(1+x) \le x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5}$ to obtain

$$A_k \;\le\; \frac{2}{3}\left(1 + \frac{2k^2 + 57k + 12}{120k^4 - 60k^3 + 40k^2 - 30k + 24}\right) \;=:\; \widehat{A}_k. \tag{10}$$

For B_k we first use $\ln(1+x) \ge x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}$ to get

$$\begin{aligned} \ln\left(\frac{k+1/2}{k(k+1)\ln(1+1/k)}\right) \;&\le\; \ln\left(1 + \frac{2k^2 - k + 3}{12k^4 + 6k^3 - 2k^2 + k - 3}\right) \\ &\le\; \frac{2k^2 - k + 3}{12k^4 + 6k^3 - 2k^2 + k - 3}, \end{aligned}$$

then we use $\ln(1+x) \ge x - \frac{x^2}{2}$ to get

$$\frac{2k+1}{\ln(1+1/k)} \;\le\; \frac{2k^2(2k+1)}{2k-1}.$$

Multiplication of both estimates yields

$$B_k \;\le\; \frac{1}{3}\left(1 + \frac{40k^3 + 14k^2 + 7k - 3}{24k^5 - 10k^3 + 4k^2 - 7k + 3}\right) \;=:\; \widehat{B}_k. \tag{11}$$

Since $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6 < \infty$, the bounds (10) and (11) on A_k and B_k, respectively, already imply that

$$\|I - T_n^*\|\,(n+1) \;=\; \frac{\sum_{k=0}^{n}(A_k + B_k)}{n+1} \;=\; 1 + O(n^{-1}).$$

To show the remaining part of the theorem, we numerically check that

$$\max_{0 \le n \le 9} \|I - T_n^*\|\,(n+1) = 2\,\|I - T_1^*\| = 1.00655\ldots$$

Table 1 The values $\|I - T_n^*\|\,(n+1)$ for $n = 2^\tau - 1$

τ	n	$\|I - T_n^*\|(n+1)$
1	1	1.00655074916
2	3	1.00498269147
3	7	1.00292347712
4	15	1.00157012523
5	31	1.00081218181
6	63	1.00041287211
7	127	1.00020813145
8	255	1.00010448957
9	511	1.00005235075
10	1023	1.00002620186
11	2047	1.00001310755
12	4095	1.00000655543
13	8191	1.00000327813
14	16383	1.00000163916
15	32767	1.00000081961

and $\widehat{A}_{10} + \widehat{B}_{10} < 1.00655\ldots$. Since $\widehat{A}_k$ and $\widehat{B}_k$ are decreasing functions of k for $k \geq 10$, we have $\widehat{A}_k + \widehat{B}_k < 1.00655\ldots$ for $k \geq 10$. Therefore,

$$\begin{aligned}
\|I - T_n^*\|\,(n+1) &\leq \frac{\sum_{k=0}^{9}(A_k + B_k) + \sum_{k=10}^{n}(\widehat{A}_k + \widehat{B}_k)}{n+1} \\
&< \frac{10 \cdot (1.00655\ldots) + (n-9)\cdot(1.00655\ldots)}{n+1} = 1.00655\ldots
\end{aligned}$$

for any $n \geq 0$, as claimed. $\square$

Table 1 shows numerically computed errors $\|I - T_n^*\|$ for specific values of $n = 2^\tau - 1$, $1 \leq \tau \leq 15$.

Remark 1 We want to stress that the choice of $\{x_i^*\}_{i=1}^n$ is very close to, but formally *not* optimal. Indeed, in the simplest case of $n = 1$, we have $x_1^* = \ln(4) \cong 1.3863$ and $\|I - T_1^*\| \cong 0.50328$, while the optimal choice is $x_1^{**} \cong 1.3568$ for which $\|I - T_1^{**}\| \cong 0.50316$.

We now estimate

$$\|T_{2n+1}^* - T_n^*\| \qquad \text{for} \quad n = 0, 1, 2, \ldots.$$

This will be much needed in the next section for the error analysis of Smolyak's algorithm for multivariate integration. Note that the values $|T_{2n+1}(f) - T_n(f)|$ are often used in constructions of various stopping criteria, e.g., in automatic integration schemes. Since

$$
\begin{aligned}
T^*_{2n+1}(f) - T^*_n(f) &= \big(I(f) - T^*_n(f)\big) - \big(I(f) - T^*_{2n+1}(f)\big)\\
&= \int_0^\infty \big(K^*_{2n+1}(t) - K^*_n(t)\big)\cdot f'(t)\,\mathrm{d}t,
\end{aligned}
$$

where K^*_m are given by (5) for the points $\{x^*_i\}_{i=1}^m$, the corresponding Peano kernel equals

$$
\overline{K}^*_n = K^*_{2n+1} - K^*_n.
$$

Using (6) and (7) we obtain that $\overline{K}^*_n$ is piecewise constant. Specifically, for $x^*_{i-1} \le t < x^*_{i-1/2}$, $1 \le i \le n$, we have

$$
\begin{aligned}
\overline{K}^*_n(t) &= \frac{\mathrm{e}^{-x^*_{i-1/2}} - \mathrm{e}^{-x^*_{i-1}}}{x^*_{i-1/2} - x^*_{i-1}} - \frac{\mathrm{e}^{-x^*_i} - \mathrm{e}^{-x^*_{i-1}}}{x^*_i - x^*_{i-1}}\\
&= -\left(\frac{x^*_i - x^*_{i-1/2}}{x^*_i - x^*_{i-1}}\right)\left(\frac{\mathrm{e}^{-x^*_i} - \mathrm{e}^{-x^*_{i-1/2}}}{x^*_i - x^*_{i-1/2}} - \frac{\mathrm{e}^{-x^*_{i-1/2}} - \mathrm{e}^{-x^*_{i-1}}}{x^*_{i-1/2} - x^*_{i-1}}\right)\\
&= -\left(x^*_i - x^*_{i-1/2}\right)\cdot D_2\left(x^*_{i-1}, x^*_{i-1/2}, x^*_i\right),
\end{aligned}
$$

where $D_2(x^*_{i-1}, x^*_{i-1/2}, x^*_i)$ denotes the second order divided difference for the weight function $\rho(x) = \mathrm{e}^{-x}$. Similarly, for $x^*_{i-1/2} \le t \le x^*_i$, $1 \le i \le n$,

$$
\overline{K}^*_n(t) = (x^*_{i-1/2} - x^*_{i-1})\cdot D_2(x^*_{i-1}, x^*_{i-1/2}, x^*_i).
$$

In the remaining cases, for $x^*_n \le x < x^*_{n+1/2}$ we have

$$
\overline{K}^*_n(t) = \frac{\mathrm{e}^{-x^*_{n+1/2}} - \mathrm{e}^{-x^*_n}}{x^*_{n+1/2} - x^*_n} = D_1(x^*_n, x^*_{n+1/2}),
$$

where $D_1(x^*_n, x^*_{n+1/2})$ is the first order divided difference for e^{-x}, and for $x^*_{n+1/2} \le t < \infty$ we have $\overline{K}^*_n(t) = 0$. Finally,

$$
\begin{aligned}
\|T^*_{2n+1} - T^*_n\| &= \int_0^\infty |\overline{K}^*_n(t)|\,\mathrm{d}t\\
&= \left(x^*_{n+1/2} - x^*_n\right) D_1\left(x^*_n, x^*_{n+1/2}\right)\\
&\quad +2\sum_{i=1}^n \left(x^*_i - x^*_{i-1/2}\right)\left(x_{i-1/2} - x^*_{i-1}\right) D_2\left(x^*_{i-1}, x^*_{i-1/2}, x^*_i\right).
\end{aligned}
$$

Plugging in the values of x^*_i given by (9), we can produce more explicit formulas; namely,

$$
\|T^*_{2n+1} - T^*_n\| = \frac{1}{(n+1)^2}\left(\frac{3}{4} + \sum_{k=1}^n \psi(k)\right), \tag{12}
$$

where

$$\psi(k) \;=\; \frac{\left(2k+\frac{1}{2}\right)\cdot\ln\left(1+\frac{1}{4k(k+1)}\right)+\ln\left(1+\frac{1}{2k}\right)}{\ln\left(1+\frac{1}{k}\right)}. \tag{13}$$

Theorem 2 *For any n we have*

$$1-\frac{1}{4(n+1)} \;<\; \|T^*_{2n+1}-T^*_n\|\,(n+1) \;<\; 1.$$

Proof We first note that $\psi(k)>1$. Indeed, by (13), this is equivalent to

$$\left(2k+\frac{1}{2}\right)\cdot\ln\left(1+\frac{1}{4k(k+1)}\right)>\ln\left(1+\frac{1}{2k+1}\right),$$

which can be verified by using $\ln(1+x)\geq\frac{1}{x}-\frac{1}{2x^2}$ for the log on the left hand side, and $\ln(1+x)\leq\frac{1}{x}-\frac{1}{2x^2}+\frac{1}{3x^3}$ for the log on the right hand side. Hence

$$\|T^*_{2n+1}-T^*_n\|\,(n+1) \;=\; \frac{3/4+\sum_{k=1}^{n}\psi(k)}{n+1} \;>\; \frac{n+3/4}{n+1} \;=\; 1-\frac{1}{4(n+1)}.$$

To show the remaining inequality, we use

$$\ln\left(1+\frac{1}{4k(k+1)}\right) \;\leq\; \frac{1}{4k(k+1)},$$

$$\ln\left(1+\frac{1}{2k}\right) \;\leq\; \frac{1}{2k}-\frac{1}{2(2k)^2}+\frac{1}{3(2k)^3} \;=\; \frac{12k^2-3k+1}{24k^3}$$

and

$$\ln\left(1+\frac{1}{k}\right) \;\geq\; \frac{1}{k}-\frac{1}{2k^2} \;=\; \frac{2k-1}{2k^2}$$

to get

$$\psi(k)\leq 1+\frac{10k+1}{24k^3+12k^2-12k}\leq 1+\frac{5}{12k^2},$$

where the last inequality is for $k\geq 2$.

For $\ell\geq 1$, let t_ℓ be defined by the equation

$$\max_{0\leq i\leq\ell}\|T^*_{2i+1}-T^*_i\|\,(i+1) \;=\; 1-t_\ell.$$

Table 2 The values $\|T_{2n+1}^* - T_n^*\|\,(n+1)$ for $n = 2^{\tau-1} - 1$

τ	$2n+1$	$\|T_{2n+1} - T_n^*\|(n+1)$
1	1	0.75000000000
2	3	0.87988750216
3	7	0.94122230503
4	15	0.97093512229
5	31	0.98554884296
6	63	0.99279476035
7	127	0.99640246605
8	225	0.99820250457
9	511	0.99910157017
10	1023	0.99955086456
11	2047	0.99977545214
12	4095	0.99988773104
13	8191	0.99994386676
14	16383	0.99997193369
15	32767	0.99998596692

Then for any $n \geq \ell + 1$

$$\begin{aligned}
\|T_{2n+1}^* - T_n^*\|\,(n+1) &= \frac{1}{n+1}\left(\frac{3}{4} + \sum_{k=1}^{\ell}\psi(k) + \sum_{k=\ell+1}^{n}\psi(k)\right) \\
&< \frac{1}{n+1}\left(\|T_{2\ell+1}^* - T_\ell^*\|(\ell+1)^2 + (n-\ell) + \frac{5}{12}\sum_{k=\ell+1}^{\infty}\frac{1}{k^2}\right) \\
&< \frac{1}{n+1}\left((1-t_\ell)(\ell+1) + (n-\ell) + \frac{5}{12\ell}\right) \\
&= 1 - \frac{1}{n+1}\left(t_\ell(\ell+1) - \frac{5}{12\ell}\right).
\end{aligned}$$

Hence the inequality $\|T_{2n+1}^* - T_n^*\|(n+1) < 1$ holds for all $n \geq \ell + 1$ if

$$t_\ell \geq \frac{5}{12\ell(\ell+1)}.$$

One can numerically check that this is true for $\ell = 2$. Since $t_2 > 0$, this completes the proof. □

Table 2 shows $\|T_{2n+1}^* - T_n^*\|$ for specific values of $n = 2^{\tau-1} - 1$, $1 \leq \tau \leq 15$.

3 Smolyak's construction for d-variate integrals

We now pass to multivariate integration. We consider the space F_d of functions $f : [0,\infty)^d \to \mathbb{R}$ that vanish at $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)$ if $x_i = 0$ for at least one $i \in \{1,\ldots,d\}$, are (locally) absolutely continuous with respect to each variable, and the norm in F_d is

$$\|f\|_{F_d} := \operatorname*{ess\,sup}_{\boldsymbol{x}\in\mathbb{R}_+^d} \left| f^{(1,\ldots,1)}(\boldsymbol{x}) \right| < \infty, \quad \text{where} \quad f^{(1,\ldots,1)} := \prod_{i=1}^{d} \frac{\partial}{\partial x_i} f.$$

Obviously, $F_1 = F$. The space F_d is the completion of

$$\operatorname{span}\left\{ f \ : \ f(\boldsymbol{x}) = \prod_{i=1}^{d} f_i(x_i) \quad \text{where} \quad f_i \in F \right\}$$

with respect to the norm $\|\cdot\|_{F_d}$. In other words, F_d is the completion of the *algebraic* d-fold tensor product $\bigotimes_{i=1}^{d} F = \underbrace{F \otimes \cdots \otimes F}_{d}$.

Since

$$f(\boldsymbol{x}) = \int_{\mathbb{R}_+} \cdots \int_{\mathbb{R}_+} f^{(1,\ldots,1)}(\boldsymbol{t}) \cdot \prod_{i=1}^{d} (x_i - t_i)_+^0 \, \mathrm{d}t_d \cdots \mathrm{d}t_1, \tag{14}$$

we have that the integration functional

$$I_d(f) := \int_{\mathbb{R}_+^d} f(\boldsymbol{x}) \cdot \exp\left(-\sum_{i=1}^{d} x_i \right) \mathrm{d}\boldsymbol{x}$$

is well defined, $I_1 = I$, and the operator norm $\|I_d\| = \|I_1\|^d = 1$.

Consider now continuous linear functionals on F of the form $L_i(f) = \sum_{k=1}^{n_i} c_{i,k}\, f(x_{i,k})$ for $i = 1,\ldots,d$. An example are the quadratures $T_{n_i}^*$. Since the algebraic tensor product $\bigotimes_{i=1}^{d} F$ is dense in F_d, the tensor product

$$\bigotimes_{i=1}^{d} L_i \ : \ F_d \to \mathbb{R}$$

defined by

$$\left(\bigotimes_{i=1}^{d} L_i \right)(f) = \prod_{i=1}^{d} L_i(f_i) \qquad \text{if} \qquad f(\boldsymbol{x}) = \prod_{i=1}^{d} f_i(x_i)$$

can be uniquely extended to a continuous linear functional on the whole F_d. Moreover, equation (14) implies that

$$\left\| \bigotimes_{i=1}^{d} L_i \right\| = \prod_{i=1}^{d} \|L_i\|.$$

We will use Smolyak's construction [17] together with the results of [24] to produce cubatures for approximating the integrals $I_d(f)$ for $f \in F_d$. That is, let τ be an integer greater than or equal to d, and let

$$Q(d,\tau) := \left\{ \boldsymbol{i} \in \mathbb{N}^d \ : \ \|\boldsymbol{i}\|_{\ell_1} \le \tau \right\}$$

be the set of $\boldsymbol{i} = [i_1, \dots, i_d]$ such that $i_k \ge 1$ and $\|\boldsymbol{i}\|_{\ell_1} = \sum_{k=1}^d i_k \le \tau$. The quadrature is given as

$$A_{d,\tau}(f) := \sum_{\boldsymbol{i}\in Q(d,\tau)} \bigotimes_{k=1}^{d} \Delta_{i_k} \quad \text{with} \quad \Delta_j := T^*_{2^j-1} - T^*_{2^{j-1}-1}. \tag{15}$$

It was shown in [24, Lemma 2] that, in generic case, the worst case error of $A_{d,\tau}$ defined as

$$\|I_d - A_{d,\tau}\| = \sup_{\|f\|_{F_d}\le 1} |I_d(f) - A_{d,\tau}(f)|,$$

depends on $C_0 = \|I\|$,

$$C_1 := \sup_{j\ge 1} \|I - T^*_{2^j-1}\| \cdot 2^j \qquad \text{and} \qquad C_2 = \sup_{j\ge 1} \|\Delta_j\| \cdot 2^j.$$

Specifically, letting $C = \max(C_0/2, C_2)$ we have

$$\|I_d - A_{d,\tau}\| \le C_1 \left(\frac{C}{2}\right)^{d-1} 2^{-(\tau-d+1)} \binom{\tau}{d-1}. \tag{16}$$

Actually, (16) was derived in [24] under the assumption that F, and consequently F_d, are Hilbert spaces; however, the same derivation carries over directly to the setting of the present paper. By (4) we have $C_0 = 1$, and by Theorems 1 and 2

$$C_1 = 2\,\|I - T^*_1\| = 1.00655\ldots \quad \text{and} \quad C_2 = 2,$$

so that the critical factor $(C/2)^{d-1}$ in (16) is just 1.

Corollary 1 *For any d and $\tau \ge d$, the worst case error of $A_{d,\tau}$ is bounded as*

$$\sup_{\|f\|_{F_d}\le 1} |I_d(f) - A_{d,\tau}(f)| = \|I_d - A_{d,\tau}\| \le C_1 \cdot 2^{-(\tau-d+1)} \cdot \binom{\tau}{d-1}. \tag{17}$$

Since the algorithm $A_{d,\tau}$ uses *nested* information, [24, Lemma 7] states that it requires exactly

$$n(d,\tau) = 2^{\tau-d+1}\binom{\tau-1}{d-1}$$

function evaluations.

Remark 2 In the next section, we will use the above results together with some results of [22]. For this reason we now restate the error bound (17) as follows. Denote

$$n+1 = 2^{\tau-d+1}.$$

Then $\tau = d - 1 + \log_2(n+1) = d - 1 + \ln(n+1)/\ln 2$,

$$\binom{\tau}{d-1} < \left(\frac{\mathrm{e}\cdot\tau}{d-1}\right)^{d-1}$$

and the error of $A_{d,\tau}$ is bounded by

$$\begin{aligned} &\mathrm{e}^{d-1}\left(\frac{C_1}{n+1}\right)\left(1+\frac{\ln(n+1)}{(d-1)\,\ln 2}\right)^{d-1} \\ &< \left(\frac{\mathrm{e}}{\ln 2}\right)^{d-1}\ \frac{C_1}{n+1}\ \left(1+\frac{\ln(n+1)}{d-1}\right)^{d-1}. \end{aligned}$$

This means that the bound (18) in [22] holds with

$$c = C_1, \quad C = \frac{\mathrm{e}}{\ln 2}, \quad \alpha_{\mathrm{std}} = 2, \quad \text{and} \quad \alpha_1 = \alpha_2 = 1.$$

4 Integration of ∞-variate functions

In this section we will use the results of the previous sections and results of [22].

4.1 Space $\mathcal{F}_{\gamma,q}$

We denote by $\boldsymbol{x} = (x_1, x_2, \dots)$ an infinite sequence (referred to as a *point*) of nonnegative reals x_i, and by $\mathbb{R}_+^{\mathbb{N}}$ the set of such points, we will also use $\mathfrak{u}$ to denote finite subsets of $\mathbb{N}$. For nonempty $\mathfrak{u}$, we will sometimes write

$$\mathfrak{u} = \{u_1, \dots, u_d\} \quad \text{with} \quad d = |\mathfrak{u}| \quad \text{and} \quad u_1 < \cdots < u_d.$$

For $\mathfrak{u} \neq \emptyset$, let $F_{\mathfrak{u}}$ denote the space like the space F_d with the only difference being that the variables of functions from $F_{\mathfrak{u}}$ are $x_{u_1}, x_{u_2}, \dots, x_{u_d}$. For $\mathfrak{u} = \emptyset$, $F_{\mathfrak{u}}$ is the space of constant functions with the natural norm, $\|c\|_{F_\emptyset} = |c|$.

Then the space $\mathcal{F}_{\gamma,q}$ is the completion of the span of functions in $\bigcup_{\mathfrak{u}\subset\mathbb{N}} F_{\mathfrak{u}}$ with respect to the norm

$$\|f\|_{\mathcal{F}_{\gamma,q}} := \left(\sum_{\mathfrak{u}\subset\mathbb{N}} \left(\|f_{\mathfrak{u}}\|_{F_{\mathfrak{u}}}/\gamma_{\mathfrak{u}}\right)^q\right)^{1/q} \quad \text{for} \quad f = \sum_{\mathfrak{u}\subset\mathbb{N}} f_{\mathfrak{u}} \text{ with } f_{\mathfrak{u}} \in F_{\mathfrak{u}}.$$

Here $q \in [1, \infty]$ and $\gamma = \{\gamma_{\mathfrak{u}}\}_{\mathfrak{u}\subset\mathbb{N}}$ is a family of nonnegative numbers, called *weights*. The role of weights has been explained in many papers; they quantify the importance of groups of variables listed in $\mathfrak{u}$'s. In particular, if $\gamma_{\mathfrak{u}} = 0$ then $f_{\mathfrak{u}} = 0$. Clearly, any $f \in \mathcal{F}_{\gamma,q}$ has the unique decomposition

$$f = \sum_{\mathfrak{u}\subset\mathbb{N}} f_{\mathfrak{u}} \quad \text{with} \quad f_{\mathfrak{u}} \in F_{\mathfrak{u}}.$$

Moreover, $\|f\|_{\mathcal{F}_{\gamma,q}} = \left[(|f_\emptyset|/\gamma_\emptyset)^q + \sum_{\mathfrak{u}\neq\emptyset}\left(\|f_{\mathfrak{u}}^{(1,\dots 1)}\|_\infty/\gamma_{\mathfrak{u}}\right)^q\right]^{1/q}$, and for $q = \infty$ we have $\|f\|_{\mathcal{F}_{\gamma,\infty}} = \sup_{\mathfrak{u}\subset\mathbb{N}} \|f_{\mathfrak{u}}^{(1,\dots,1)}\|_\infty/\gamma_{\mathfrak{u}}$. Although some of the results below hold for general weights, we restrict the attention to an important class of POD weights, see, (2), that were introduced in [10].

4.2 Integration problem

We are interested in efficient algorithms for approximating

$$\mathcal{I}_\infty(f) := \lim_{d\to\infty} \int_{\mathbb{R}_+^d} f(x_1, \ldots, x_d, 0, 0, \ldots) \cdot \exp\left(-\sum_{i=1}^d x_i\right) \mathrm{d}\boldsymbol{x}.$$

The functional $\mathcal{I}_\infty$ is a well defined and continuous functional on $\mathcal{F}_{\gamma,q}$ iff

$$\beta_2 > \frac{1}{q^*}, \quad \text{where} \quad \frac{1}{q} + \frac{1}{q^*} = 1. \tag{18}$$

This is because $\|\mathcal{I}_\infty\| = \left(\sum_{\mathfrak{u}\subset\mathbb{N}} \gamma_{\mathfrak{u}}^{q^*}\right)^{1/q^*}$ and

$$\sum_{\mathfrak{u}\subset\mathbb{N}} \gamma_{\mathfrak{u}}^{q^*} < \infty \quad \text{iff} \quad \sum_{j=1}^\infty j^{-\beta_2 q^*} < \infty.$$

Of course, for $q^* = \infty$, $\|\mathcal{I}_\infty\| = \max_{\mathfrak{u}\subset\mathbb{N}} \gamma_{\mathfrak{u}}$ and (18) reduces to $\beta_2 > 0$. From now on we assume (18).

Remark 3 One could consider approximating integrals with the density of standard exponential probability e^{-x_i} replaced by $\rho_i(x_i) = \lambda_i^{-1} \cdot \mathrm{e}^{-x_i/\lambda_i}$. Since

$$\lambda_i^{-1} \int_0^\infty f(x_i)\, \mathrm{e}^{-x_i/\lambda_i}\, \mathrm{d}x_i = \int_0^\infty g(x_i)\, \mathrm{e}^{-x_i}\, \mathrm{d}x \quad \text{for} \quad g(x_i) = f(\lambda_i x_i),$$

the problem with λ_i different from 1 is equivalent to the problem considered in this paper; however, with the weights $\gamma_{\mathfrak{u}}$ replaced by $\gamma_{\mathfrak{u}} \prod_{i\in\mathfrak{u}} \lambda_i$. In particular, the integral operator with $\lambda_i \neq 1$ is continuous iff

$$\left(\sum_{\mathfrak{u}\subset\mathbb{N}} \gamma_{\mathfrak{u}}^{q^*} \prod_{i\in\mathfrak{u}} \lambda_i^{q^*}\right)^{1/q^*} < \infty.$$

As mentioned in the Introduction, we assume that each function $f \in \mathcal{F}_{\gamma,q}$ can be evaluated at points $\boldsymbol{x}$ that have only a finite number $d(\boldsymbol{x})$ of nonzero coordinates, and the cost of such an evaluation is

$$\$(d(\boldsymbol{x})).$$

Here $\$: \mathbb{N} \cup \{0\} \to \mathbb{R}_+$ is a given monotonically increasing function, called *cost function*. Then the cost of an algorithm is the total cost of the samples $f(\boldsymbol{x}_i)$ used by it. For instance, the cost of $A(f) = \sum_{i=1}^n a_i\, f(\boldsymbol{x}_i)$ is equal to $\sum_{i=1}^n \$(d(\boldsymbol{x}_i))$.

4.3 Multivariate decomposition method

We use the following notation:

$$\bar{\gamma}_{\mathfrak{u}} = C_0^{|\mathfrak{u}|}\, \gamma_{\mathfrak{u}} \quad \text{and} \quad \widehat{\gamma}_{\mathfrak{u}} = C_1 \left(\frac{\mathrm{e}}{\ln 2}\right)^{|\mathfrak{u}|-1} \gamma_{\mathfrak{u}}.$$

In the first step of MDM, we construct a set $\mathfrak{U}(\varepsilon) = \mathfrak{U}(\varepsilon, q)$ that is possibly small and such that $\mathcal{I}(f_{\mathfrak{u}})$ can be approximated by zero for $\mathfrak{u} \notin \mathfrak{U}(\varepsilon)$. More precisely, for $q = 1$, we take

$$\mathfrak{U}(\varepsilon) := \left\{ \mathfrak{u} \subset \mathbb{N} \, : \, \bar{\gamma}_{\mathfrak{u}} > \frac{\varepsilon}{2} \right\}.$$

For $q > 1$, take any κ satisfying

$$0 < \kappa < \beta - \frac{1}{q^*},$$

and

$$\mathfrak{U}(\varepsilon) := \left\{ \mathfrak{u} \subset \mathbb{N} \, : \, \bar{\gamma}_{\mathfrak{u}} \geq \left(\frac{\varepsilon}{2 \cdot [L(q^*, \kappa)]^{1/q^*}} \right)^{(\kappa + 1/q^*)/\kappa} \right\},$$

where

$$L(q^*, \kappa) := \sum_{\mathfrak{v} \in \mathfrak{U}_\gamma} \bar{\gamma}_{\mathfrak{u}}^{q^*/(\kappa \cdot q^* + 1)}.$$

Since $q^*/(\kappa \cdot q^* + 1) > 1/\beta$, $L(q^*, \kappa)$ is well defined. Of course, for $q > 1$, $\mathfrak{U}(\varepsilon)$ depends also on κ, $\mathfrak{U}(\varepsilon) = \mathfrak{U}(\varepsilon, q, \kappa)$.

We have from [22, Thm. 2 and Prop. 3] that $\sum_{\mathfrak{u} \notin \mathfrak{U}(\varepsilon)} |\mathcal{I}(f_{\mathfrak{u}})| \leq \|f\|_{\mathcal{F}_{\gamma,q}} \varepsilon/2$. Moreover, $|\mathfrak{U}(\varepsilon)| = O\left(\varepsilon^{-1/\kappa}\right)$ for any $\kappa \in (0, \beta_2 - 1/q^*)$. As in [15], one can verify that

$$d(\varepsilon) = d(\varepsilon, q) := \max_{\mathfrak{u} \in \mathfrak{U}(\varepsilon)} |\mathfrak{u}| \leq c \cdot \frac{\ln(1/\varepsilon)}{\ln(\ln(\mathrm{e}/\varepsilon))} \tag{19}$$

for some positive constant c.

In the second step of MDM, each $\mathcal{I}(f_{\mathfrak{u}})$ for $\mathfrak{u} \in \mathfrak{U}(\varepsilon)$ is approximated separately. More precisely,

$$\mathcal{A}_\varepsilon(f) = \sum_{\mathfrak{u} \in \mathfrak{U}(\varepsilon)} A_{\mathfrak{u}, \tau_{\mathfrak{u}}}(f_{\mathfrak{u}}),$$

where $A_{\mathfrak{u},\tau_{\mathfrak{u}}}$ are Smolyak's algorithms (15) adopted to the spaces $F_{\mathfrak{u}}$, i.e., instead of $x_1, \ldots, x_d$ they work on variables $x_{u_1} \ldots, x_{u_q}$. For $\mathfrak{u} = \emptyset$ we have $A_{\emptyset,1} = f(\mathbf{0})$.

Remark 4 Although we do not know $f_{\mathfrak{u}}$, their samples can be obtained from $2^{|\mathfrak{u}|}$ values of the function f, see [12] Then the cost of such sampling of $f_{\mathfrak{u}}$ is bounded by $2^{|\mathfrak{u}|} \cdot \$(|\mathfrak{u}|) \leq 2^{d(\varepsilon)} \cdot \$(d(\varepsilon))$, where $d(\varepsilon)$ is as in (19). Hence it is relatively small even if \$ is an exponential function.

The numbers $\tau_{\mathfrak{u}}$ depend on an additional parameter δ. This is why the algorithm $\mathcal{A}_\varepsilon = \mathcal{A}_{\varepsilon,\delta}$ also depends on δ. Choose

$$\delta \geq \frac{1}{2},$$

and define

$$h_{\mathfrak{u}} := \left(\frac{2}{\varepsilon}\right)^{2 \cdot \delta} \cdot \widehat{\gamma}_{\mathfrak{u}}^{2 \cdot \delta}$$

for $q = 1$ and

$$h_{\mathfrak{u}} := \left(\frac{2}{\varepsilon}\right)^{2\cdot\delta} \widehat{\gamma}_{\mathfrak{u}}^{2\cdot\delta\cdot q^*/(q^*+2\cdot\delta)} \left[\sum_{\mathfrak{v}\in\mathfrak{U}(\varepsilon)} \widehat{\gamma}_{\mathfrak{v}}^{2\cdot\delta\cdot q^*/(q^*+2\cdot\delta)}\right]^{2\cdot\delta/q^*}$$

for $q > 1$. Then

$$\tau_{\mathfrak{u}} := |\mathfrak{u}| + \max\left(0\,,\ \log_2(\lfloor h_{\mathfrak{u}}\rfloor + 1) - 1\right).$$

Clearly $\lfloor h_{\mathfrak{u}}\rfloor + 1 \leq 2^{\tau_{\mathfrak{u}}-|\mathfrak{u}|+1} < 2\cdot(\lfloor h_{\mathfrak{u}}\rfloor + 1)$. Following [22, Thm. 9, Prop. 10, and Prop. 11] and Remark 2 we have

$$e(\mathcal{A}_{\varepsilon,\delta}; \mathcal{F}_{\gamma,q}) := \sup_{\|f\|_{\mathcal{F}_{\gamma,q}\leq 1}} |\mathcal{I}_\infty(f) - \mathcal{A}_{\varepsilon,\delta}(f)| \leq \varepsilon\cdot B(\varepsilon, q),$$

where

$$B(\varepsilon, q) = \max_{\mathfrak{u}\in\mathfrak{U}(\varepsilon)} \left(1 + \frac{\ln(\lfloor h_{\mathfrak{u}}\rfloor + 1)}{|\mathfrak{u}| - 1}\right)^{|\mathfrak{u}|-1},$$

and

$$\mathrm{cost}(\mathcal{A}_{\varepsilon,\delta}) \leq 2^{1+d(\varepsilon)}\cdot \$(d(\varepsilon))\cdot\left(\frac{2}{\varepsilon}\right)^{2\cdot\delta}\cdot\left[\sum_{\mathfrak{u}\in\mathfrak{U}(\varepsilon)} \widehat{\gamma}_{\mathfrak{u}}^{2\cdot\delta\cdot q^*/(q^*+2\cdot\delta)}\right]^{(q^*+2\cdot\delta)/q^*}.$$

In general, if δ is too small, then the last sum may converge to infinity with ε tending to zero. Hence there could be a trade-off between the exponent $2\,\delta$ of $1/\varepsilon$ and the sum. However, if, in addition to $\delta \geq 1/2$, we have $2\,\delta > 1/(\beta_2 - 1/q^*)$, then the last sum is uniformly bounded for every $\varepsilon > 0$. Moreover,

$$B(\varepsilon, q) = \varepsilon^{-O(\ln(\ln(\ln(1/\varepsilon)))/\ln(\ln(1/\varepsilon)))}$$

as $\varepsilon \to 0$. This yields the following theorem.

Theorem 3 *Suppose that $\beta_2 > 1/q^*$.*

(***i***) *If $\$(d) = \mathrm{e}^{O(d)}$, then the problem is polynomially tractable with the tractability exponent*

$$p^{\mathrm{trct}} \leq \max\left(1\,,\ \frac{1}{\beta_2 - 1/q^*}\right).$$

If, additionally, $\$(d) = \Omega(d)$, then we have equality above.

(***ii***) *If $\$(d) = \mathrm{e}^{\mathrm{e}^{O(d)}}$, then the problem is weakly tractable.*

Acknowledgments A part of the research presented in this paper was conducted when the authors were visiting the University of New South Wales. We would like to express gratitude to our hosts Frances Y. Kuo and Ian. H. Sloan.

References

1. Baldeaux, J.: Scrambled polynomial lattice rules for infinite-dimensional integration, MCQMC 2010. In: Plaskota, L., Woźniakowski, H. (eds) Springer Proceedings in Mathematics and Statistics, pp. 255–263. Springer-Verlag, Berlin (2012)

2. Baldeaux, J., Gnewuch, M.: Optimal randomized multilevel algorithms for infinite-dimensional integration on function spaces with ANOVA-type decomposition. arXiv:1209.0882v1 (2012)
3. Creutzig, J., Dereich, S., Müller-Gronbach, T., Ritter, K.: Infinite-dimensional quadrature and approximation of distributions. Found. Comput. Math. **9**, 391–429 (2009)
4. Gnewuch, M.: Infinite-dimensional integration on weighted Hilbert spaces. Math. Comp. **81**, 2175–2205 (2012)
5. Gnewuch, M.: Lower error bounds for randomized multilevel and changing dimension algorithms. (2012, in press)
6. Gnewuch, M., Mayer, S., Ritter, K.: On weighted Hilbert spaces and integration of functions of infinitely many variables. J. Complex (2013, in press)
7. Hickernell, F.J., Müller-Gronbach, T., Niu, B., Ritter, K.: Multi-level Monte Carlo algorithms for infinite-dimensional integration on $\mathbb{R}^{\mathbb{N}}$. J. Complex. **26**, 229–254 (2010)
8. Hickernell, F.J., Wang, X.: The error bounds and tractability for quasi-Monte Carlo algorithms in infinite dimension. Math. Comp. **71**, 1641–1661 (2001)
9. Kuo, F.Y., Schwab, C., Sloan, I.H.: Quasi-Monte Carlo methods for high-dimensional integration: the standard (weighted Hilbert space) setting and beyond. ANZIAM J. 1–37 (2011)
10. Kuo, F.Y., Schwab, C., Sloan, I.H.: Quasi-Monte Carlo finite element methods for a class of partial differential equations with random coefficients, Res. Report 2011-52, Seminar for Applied Mathematics, ETH Zürich. SIAM J. Numer. Anal. (2011)
11. Kuo, F.Y., Sloan, I.H., Wasilkowski, G.W., Woźniakowski, H.: Liberating the dimension. J. Complex. **26**, 422–454 (2010). doi:10.1016/j.jco.2009.12.003
12. Kuo, F.Y., Sloan, I.H., Wasilkowski, G.W., Woźniakowski, H.: On decompositions of multivariate functions. Math. Comput. **79**, 953–966 (2010). doi:0.1090/S0025-5718-09-02319-9
13. Niu, B., Hickernell, F.J.: Monte Carlo simulation of stochastic integrals when the cost function evaluation is dimension dependent. In: Ecuyer, P.L., Owen, A.B. (eds.) Monte Carlo and Quasi-Monte Carlo Methods 2008, pp. 545–560. Springer (2008)
14. Niu, B., Hickernell, F.J., Müller-Gronbach, T., Ritter, K.: Deterministic multi-level algorithms for infinite-dimensional integration on $\mathbb{R}^{\mathbb{N}}$. J. Complex. **27**, 331–351 (2011)
15. Plaskota, L., Wasilkowski, G.W.: Tractability of infinite-dimensional integration in the worst case and randomized settings. J. Complex. **27**, 505–518 (2011). doi:10.1016/j.com.2011.01.006
16. Plaskota, L., Wasilkowski, G.W., Woźniakowski, H.: A new algorithm and worst case complexity for Feynman-Kac path integration. J. Comput. Phys. **164**, 335–353 (2000)
17. Smolyak, S.A.: Quadrature and interpolation formulas for tensor products of certain classes of functions. Dokl. Acad. Nauk SSSR **4**, 240–243 (1963)
18. Traub, J.F., Wasilkowski, G.W., Woźniakowski, H.: Information-Based Complexity. Academic, New York (1988)
19. Wasilkowski, G.W.: Liberating the dimension for function approximation and integration, MCQMC 2010. In: Plaskota, L., Woźniakowski, H. (eds.) Springer Proceedings in Mathematics and Statistics, pp. 211–231. Springer-Verlag, Berlin (2012). doi:10.1007/978-3-642-27440-4-9
20. Wasilkowski, G.W.: Liberating the dimension for L_2-approximation. J. Complex. **28**, 304–319 (2012). doi:10.1016/j.jco.2011.12.002
21. Wasilkowski, G.W.: Average case tractability of approximating ∞-variate functions. Math. Comp.electronically published (2013)
22. Wasilkowski, G.W.: On tractability of linear tensor product problems for ∞-variate classes of functions. J. Complex. **29**, 351–369 (2013). doi:10.1016/j.jco.2013.04.008
23. Wasilkowski, G.W.: Tractability of approximation of ∞-variate functions with bounded mixed partial derivatives, submitted
24. Wasilkowski, G.W., Woźniakowski, H.: Explicit cost bounds for multivariate tensor product problems. J. Complex. **11**, 1–56 (1995)
25. Wasilkowski, G.W., Woźniakowski, H.: On tractability of path integration. J. Math. Phys. **37**, 2071–2088 (1996)
26. Wasilkowski, G.W., Woźniakowski, H.: Liberating the dimension for function approximation. J. Complex. **27**, 86–110 (2011). doi:10.10.16/jco.2010.08.004
27. Wasilkowski, G.W., Woźniakowski, H.: Liberating the dimension for function approximation: standard information. J. Complex. **27**, 417–440 (2011). doi:10.10.16/j.jco.2011.02.002